BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 8: e136016
doi: 10.3897/biss.8.136016

OPEN ACCESS

Conference Abstract

A Taxonomic Concept Mapping Service for Taxonomic Information Aggregators

Andreas Müller‡, Eckhard von Raab-Straube‡, Walter G. Berendsohn‡

‡ Botanic Garden and Botanical Museum, Freie Universität, Berlin, Germany

Corresponding author: Andreas Müller (a.mueller@bo.berlin)

Received: 31 Aug 2024 | Published: 02 Sep 2024

Citation: Müller A, von Raab-Straube E, Berendsohn WG (2024) A Taxonomic Concept Mapping Service for Taxonomic Information Aggregators. Biodiversity Information Science and Standards 8: e136016. https://doi.org/10.3897/biss.8.136016

Abstract

Linking local taxon-related information to global (or larger area) aggregators of taxonomic information is in ever growing demand. However, linking local taxon-related data to global information systems faces various obstacles. Apart from technical problems, there are specific taxonomic issues to solve. Not only do taxon names need to be matched if written in different ways (e.g., *Abies alba* Mill. and *Abies alba* Miller), but also taxonomic concepts have to be compared and mapped. A taxonomic concept is here understood to circumscribe the taxon, i.e., it determines to which taxon an individual organism belongs. The international rules of nomenclature then determine what name has to be applied to the taxon. Unlike names, which can be mapped to each other objectively, concepts often can not be mapped one-to-one. Even when bearing the same accepted name, concepts may be wider or narrower in the local list than in the global list. In addition, taxa may be fully missing in a local list while they exist in the aggregator, simply because they are not considered to be independent taxa in the first place or they do not occur in a list with a regional geographic or thematic scope.

Different approaches exist to define taxon concepts (Geoffroy and Berendsohn 2003). One is to use textual circumscriptions defining the characteristics of a taxon. Frequently, however, circumsciptions do not exist in the given context or they might be difficult to use for automated taxon concept comparison. Therefore, another—and in this context more promising—approach is to compare taxon names used in the synonymies of the given

© Müller A et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

taxa including the accepted names. This approach takes advantage of the fact that taxon names are placeholders for type specimens and, therefore, the sum of all names defines a group of type specimens. In a second step of deduction, all the characteristics of these specimens can then be used to define the characteristics of the taxa they belong to (Pyle et al. 2022). Depending on the quality and completeness of the synonymies and taxonomic backbones involved, this fact can be used to distinguish the relationships between taxa relatively well, even without describing the individual characters explicitly.

As the described approach is still complex and depends on certain prerequisites (e.g., the presence of synonyms in the local lists), some global aggregators, like the World Flora Online (WFO), use a simpler approach at first, by taking the accepted name in the local list, looking it up in the synonymy of the global list and linking the data related to the local taxon to the accordingly accepted taxon of the global list.

However, this approach may lead to inaccuracy and irritation among users because it does not take into account the differences in taxonomic concepts. For example, a conservationist in a given country may wonder why a taxon classified as threatened in their country is not classified as such in other countries. They may assume the cause to be environmental conditions when in reality it is merely a function of the differences in the underlying taxonomic concepts used.

As a contribution to the WFO initiative and following the concepts of a service-oriented modular architecture, we will develop a standalone service that integrates the computation of taxonomic concept relationships with a user-friendly visualization and explanation of the relationships. WFO content providers may decide to use this module by providing synonymies along with their content. The accuracy of the results will depend on the quality and completeness of the input. For example, an incomplete synonymy on the local side will increase the uncertainty of the computed taxon relationships. The same holds true for incomplete taxonomic coverage on the aggregator's side, for taxon relationships are not only computed by comparing the synonymies of the two or three taxa involved but also by checking the existence of names in the synonymy of other taxa.

Matching the local accepted names with WFO names and providing WFO identifiers (IDs) along with the content are a standard but time consuming requirement for providing content to WFO. To minimize the potential barrier created by data entry requirements for the concept relations module, it will not be obligatory to provide WFO-IDs together with the synonyms, though adding them will increase the quality of the results. The option to provide information on homotypic groups (names sharing the same type) will also be a feature. Providing homotypic groups will increase accuracy as concepts will now be defined by their included type specimens. Linking names unambiguously to their types is one essential step in computing accurate concept relationships.

A comprehensive understanding of the consequences deriving from the computed concept relationships is critical when using this module. Thus, the results will be explained and visualized in multiple ways. While visualization via set semantics or by comparison of synonymies might be useful for experts, textual explanations of the

relationships as well as their consequences are likely to be more helpful for the majority of users.

As inputs, this service will take commonly used formats for taxonomic backbone information such as the Darwin Core Archive (DwC-A), the Catalogue of Life Data Package (ColDP), and well-defined web service APIs.

The service will be developed for usage in the WFO but can be applied in other contexts as well.

Keywords

synonymy mapping, World Flora Online

Presenting author

Andreas Müller

Presented at

SPNHC-TDWG 2024

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Geoffroy M, Berendsohn W (2003) The concept problem in taxonomy: Importance, components, approaches. In: Berendsohn W (Ed.) Schriftenreihe Vegetationskunde. 39. Federal Agency for Nature Conservation, Bonn, 113 pp. URL: https://www.researchgate.net/publication/285456910_The_concept_problem_in_taxonomy_Importance_components_approaches
- Pyle R, Bailly N, Remsen D (2022) Modeling Taxon Concepts: A new approach to an old problem. Biodiversity Information Science and Standards 6 https://doi.org/10.3897/biss.6.93927